

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 13:01:31 ; Search time 3.69565 Seconds  
(without alignments)  
112.230 Million cell updates/sec

Title: US-09-801-784A-36  
Perfect score: 50  
Sequence: 1 PSAVALTSP 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

c-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	88.0	171	1 FMS1_ECOLI	P25730 escherichia
2	40	80.0	170	1 PMCI_ECOLI	P02971 escherichia
3	36	72.0	234	1 POLN_RHDV3	P27411 rabbit hemo
4	36	72.0	326	1 MTC2_CHVP1	P31118 parametium
5	36	72.0	377	1 VMAT_P12HT	P24466 human para
6	36	72.0	2344	1 POLN_RHDV	P27410 rabbit hemo
7	35	70.0	374	1 ADH7_HUMAN	P40394 homo sapien
8	35	70.0	374	1 ADH7_MOUSE	O64437 homo sapien
9	35	70.0	486	1 RTG3_YEAST	P38165 saccharomyc
10	35	70.0	845	1 KRUI_DROME	P14180 saccharomyc
11	35	70.0	963	1 CHS2_YEAST	P14180 saccharomyc
12	35	70.0	2663	1 LRBA_HUMAN	P50851 homo sapien
13	34	68.0	540	1 ANPC_HUMAN	P17342 homo sapien
14	34	68.0	736	1 DVU2_MOUSE	O60838 mus musculu
15	34	68.0	943	1 BLI4_CAEEL	P51559 caenorhabdi
16	34	68.0	2016	1 CINS_HUMAN	O14524 homo sapien
17	33.5	67.0	1291	1 PYRD_SULAC	O08358 sulfolobus
18	33	66.0	314	1 GBLP_SCHPO	O10281 schizosacch
19	33	66.0	369	1 VMAT_LPMV	P27287 la piedad-m
20	33	66.0	375	1 ADHA_UFOHA	P25405 utromastix h
21	33	66.0	382	1 VMAT_SV41	P25405 utromastix h
22	33	66.0	438	1 EFIA_DESMO	P41203 desulfuroco
23	33	66.0	547	1 FIB1_ADE40	P18047 human adeno
24	33	66.0	561	1 ATKA_ANASU	O916X2 anabaena sp
25	33	66.0	664	1 PTSA_STRMT	P12555 streptococ
26	33	66.0	680	1 DAPI_HUMAN	O15228 homo sapien
27	33	66.0	919	1 PWP2_HUMAN	O15269 homo sapien
28	33	66.0	1274	1 SRP3_BACSU	O08787 bacillus su
29	33	66.0	1301	1 AKX1_HUMAN	O9444 homo sapien
30	32	64.0	123	1 SDIS_NCCSI	P77814 nocardioid
31	32	64.0	169	1 CU21_LOCM1	P81225 locusta mig
32	32	64.0	252	1 BTG3_MOUSE	P50615 mus musculu
33	32	64.0	307	1 YL72_ARCPU	O28110 archaeoglob

34	32	64.0	326	1 Y800_PYPAB	Q0u2e8 pyrococcus
35	32	64.0	384	1 SDG3_HUMAN	O75056 homo sapien
36	32	64.0	428	1 EPIA_PYPAB	O9v077 pyrococcus
37	32	64.0	428	1 EPIA_PYPAB	O59153 pyrococcus
38	32	64.0	430	1 EPIA_PYPAB	P26751 pyrococcus
39	32	64.0	437	1 EPIA_AERRE	Q9yav0 aetopyrum p
40	32	64.0	787	1 DPOL_HPBDC	P30028 duck hepati
41	32	64.0	788	1 DPOL_HPBDC	P17192 duck hepati
42	32	64.0	788	1 DPOL_HPBDC	P17193 duck hepati
43	31	62.0	85	1 HPIS_THIRO	P00261 thiorocapsa r
44	31	62.0	194	1 YBEA_SCHPO	O14256 schizosacch
45	31	62.0	274	1 VGLF_HSV2	P13285 herpes simp

## ALIGNMENTS

RESULT 1  
FMS1\_ECOLI STANDARD; PRT; 171 AA.  
AC P25730;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE CSI fimbrial subunit A precursor (CSI pilin).  
GN CSOA OR COOA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RX MEDLINE=91357443; PubMed=1679404;  
RA Jorda B.U.A.M., van Vliet A.H.M., Willems G.A.,  
RA van der Zeijst B.A.M., van Vliet A.H.M., Willems G.A.,  
RT "Analysis of the first two genes of the CSI fimbrial operon in human  
RT enterotoxigenic Escherichia coli of serotype O139:H28."  
RL FEMS Microbiol. Lett. 64:265-270(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H16; PLASMID=pE605;  
RX MEDLINE=91034170; PubMed=1977705;  
RA Perez-Casal J., Swartley J.S., Scott J.R.;  
RT "Gene encoding the major subunit of CSI pilin of human enterotoxigenic  
RT Escherichia coli."  
RL Infect. Immun. 58:3594-3600(1990).  
RN [3]  
RP SEQUENCE OF 24-42.  
RC STRAIN=60875;  
RX MEDLINE=90036735; PubMed=2572563;  
RA Hall R.H., Maneval D.R. Jr., Collins J.H., Theibert J.L.,  
RA Levine M.W.;  
RT "Purification and analysis of colonization factor antigen 1, coli  
RT surface antigen 1, and coli surface antigen 3 fimbriae from  
RT enterotoxigenic Escherichia coli."  
RL J. Bacteriol. 171:6372-6374(1989).  
RN [4]  
RP FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING  
RP FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5  
RP MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO  
RP COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.  
RN [5]  
RP INDUCTION: CSI FIMBRIAE ARE ONLY EXPRESSED IN THE PRESENCE OF THE  
RP POSITIVE REGULATOR RNS.  
RN [6]  
RP SIMILARITY: TO THE CFA/I FIMBRIAL SUBUNIT B (CFAI).  
RN [7]  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
RP between the Swiss Institute of Bioinformatics and the EMBL outstation  
RP at the European Bioinformatics Institute. There are no restrictions on its  
RP use by non-profit institutions as long as its content is in no way  
RP modified and this statement is not removed. Usage by and for commercial  
RP entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
RP or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

```

CC -----
DR EMBL; X62879; CAA44673.1; -.
DR EMBL; M58550; AAA23596.1; -.
DR PIR; A41467; A41467.
DR PIR; S19003; S19003.
KW Fimbrin; Signal; Plasmid.
FT SIGNAL 1 23
FT CHAIN 24 171 CS1 FIMBRIAL SUBUNIT A.
FT CONFLICT 28 28 I -> E (IN REF. 3).
SQ SEQUENCE 171 AA; 17542 MW; 46E70BE7053DBE13 CRC64;

Query Match 88.0%; Score 44; DB 1; Length 171;
Best Local Similarity 80.0%; Pred. No. 0.09;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSVAITYSP 10
   |||||
Db 49 PSVAITYSP 58

RESULT 2
ID FMC1 ECOLI
AC P02971; STANDARD; PRT; 170 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1993 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CFA/I fimbrial subunit B precursor [Colonization factor antigen 1
DE subunit B] (CFA/I pilin) (CFA/I antigen).
GN CFAB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
CX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89173309; PubMed=2564374;
RA Karjalainen T.K., Evans D.G., So M., Lee C.-H.;
RT "Molecular cloning and nucleotide sequence of the colonization factor
RT antigen 1 gene of Escherichia coli.";
RL Infect. Immun. 57:1126-1130(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89330163; PubMed=2569152;
RA Hamers A.M., Pel H.J., Willshaw G.A., Kusters J.G.,
RA van der Zeijst B.A.M., Gaastra W.;
RT "The nucleotide sequence of the first two genes of the CFA/I fimbrial
RT operon of human enterotoxigenic Escherichia coli.";
RL Microb. Pathog. 6:297-309(1989).
RN [3]
RP SEQUENCE OF 24-170 FROM N.A.
RX STRAIN=O78:H11 / H10407;
RC MEDLINE=82235736; PubMed=6124420;
RA Klemm P.;
RT "Primary structure of the CFA/I fimbrial protein from human
RT enterotoxigenic Escherichia coli strains.";
RL Eur. J. Biochem. 124:339-348(1982).
RN [4]
RP SEQUENCE OF 24-42.
RX STRAIN=O78:H11 / H10407;
RC MEDLINE=90036735; PubMed=2572583;
RA Hall R.H., Maneval D.R. Jr., Collins J.H., Theibert J.L.,
RA Levine M.M.;
RT "Purification and analysis of colonization factor antigen I, coli
RT surface antigen 1, and coli surface antigen 3 fimbriae from
RT enterotoxigenic Escherichia coli.";
RL J. Bacteriol. 171:6372-6374(1989).
CC -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC -!- SUBUNIT: CFA/I FIMBRIAE ARE RATHER RIGID, THREAD-LIKE FILAMENTS OF

```

```

CC 0.5-1 MICROMETER, WITH AN APPARENT AXIAL HOLE, AND A DIAMETER OF
CC 7 NANOMETERS. A SINGLE CFA/I FIMBRIA CONSISTS OF ABOUT 100
CC IDENTICAL PROTEIN SUBUNITS.
CC -!- INDUCTION: CFA/I FIMBRIAE ARE ONLY EXPRESSED IN THE PRESENCE OF
CC THE POSITIVE REGULATOR CFAD.
CC -!- SIMILARITY: TO THE CS1 FIMBRIAL SUBUNIT A (CSOA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M55661; AAC41415.1; -.
DR PIR; A30589; YQECCL.
KW Fimbrin; Antigen; Plasmid; Signal.
FT SIGNAL 1 23
FT CHAIN 24 170 CFA/I FIMBRIAL SUBUNIT B.
FT CONFLICT 37 37 V -> A (IN REF. 2).
FT CONFLICT 76 76 D -> N (IN REF. 3).
FT CONFLICT 97 97 S -> A (IN REF. 3).
SQ SEQUENCE 170 AA; 17461 MW; 924347F07F33CDAB CRC64;

Query Match 80.0%; Score 40; DB 1; Length 170;
Best Local Similarity 80.0%; Pred. No. 0.57;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSVAITYSP 10
   |||||
Db 49 PSVAITYSP 58

RESULT 3
ID POLN RHDV3 STANDARD; PRT; 234 AA.
AC P27411;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-structural polyprotein [Contains: Coat protein] (Fragment).
OS Rabbit hemorrhagic disease virus (strain V-351) (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
CX NCBI_TaxID=11977;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92307403; PubMed=1497750;
RA Milton I.D., Vlasak R., Nowotny N., Rodak L., Carter M.J.;
RT "Genomic 3' terminal sequence comparison of three isolates of rabbit
RT haemorrhagic disease virus.";
RL FEMS Microbiol. Lett. 72:37-42(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11535; CAA77633.1; -.
DR PIR; S22134; S22134.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Polyprotein; Coat protein.
FT DOMAIN 1 1
FT NON_TER 7 234
SQ SEQUENCE 234 AA; 23777 MW; B66018654E2EE05A CRC64;

Query Match 72.0%; Score 36; DB 1; Length 234;
Best Local Similarity 60.0%; Pred. No. 5.2;

```

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSAVALTYSP 10  
|:|:|:|:|  
Db 84 PMSALTYTP 93

RESULT 4  
ID MTC2\_CHVP1 STANDARD; PRT; 326 AA.  
AC P1118; 084569;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Modification methylase CVAI1 (EC 2.1.1.72) (Adenine-specific  
methyltransferase CVAI1) (M.CVAI1).  
GN CVAI1 OR A251R.  
OS Paramyxium bursaria chlorella virus 1 (PBCV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.  
OX NCBI\_TaxID=10506;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=93065201; PubMed=1437552;  
RA Zhang Y., Nelson M., Niefeldt J.W., Burbank D.B., van Etten J.L.;  
"Characterization of Chlorella virus PBCV-1 CVAI1 restriction and  
modification system.";  
Nucleic Acids Res. 20:5351-5356 (1992).  
[2]  
REVIEWS, AND SEQUENCE FROM N.A.  
RX MEDLINE=9313167; PubMed=7831789;  
RA Lu Z., Li Y., Zhang Y., Kutish G.F., Rock D.L., van Etten J.L.;  
"Analysis of 45 kb of DNA located at the left end of the chlorella  
virus PBCV-1 genome.";  
Virology 206:339-352 (1995).  
-1- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE  
CATCG. CAUSES SPECIFIC METHYLATION ON A-2 ON BOTH STRANDS AND  
PROTECTS THE DNA FROM CLEAVAGE BY THE CVAI1 ENDONUCLEASE.  
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-  
adenosyl-L-homocysteine + DNA 6-methylaminopurine.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M86639; AAB92381.1; -;  
DR EMBL; U42580; AAC96619.1; -;  
DR PIR; S27901; S27901.  
DR PIR; S35442; S35442.  
REBASE; 3542; M.CVAI1.  
InterPro; IPR002294; D12N6\_mtfase.  
InterPro; IPR002052; N6\_Mtase.  
Pfam; PF00806; Methyltransferase.  
DR PRINTS; PR00505; D12N6MTFAS.  
DR PROSITE; PS00092; N6\_MTFASE.1.  
KM Transferase; Methyltransferase; Restriction system.  
SQ SEQUENCE 326 AA; 37492 MW; 8CC1FC75D85F0D89 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 326;  
Best Local Similarity 70.0%; Pred. No. 7.5;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PSAVALTYSP 10  
|:|:|:|:|  
Db 95 PGFVLTYS 104

RESULT 5  
WMT\_P12HT STANDARD; PRT; 377 AA.  
ID WMT\_P12HT

AC P24266;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Matrix protein (membrane protein).  
GN M.  
OS Human parainfluenza 2 virus (strain Toshiba) (PIV-2).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.  
OX NCBI\_TaxID=11214;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91049454; PubMed=2173264;  
RA Kawano M., Bando H., Ohgimoto S., Okamoto K., Kondo K., Tsurudome M.,  
RA Nishio M., Ito Y.;  
RT "Complete nucleotide sequence of the matrix gene of human  
parainfluenza type 2 virus and expression of the M protein in  
bacteria.";  
Virology 179:857-861 (1990).  
RL -1- FUNCTION: THE M PROTEIN HAS A CRUCIAL ROLE IN VIRUS ASSEMBLY  
AND INTERACTS WITH THE RNP COMPLEX AS WELL AS WITH THE VIRAL  
MEMBRANE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M62734; AAA46862.1; -;  
DR EMBL; X57559; CAA40785.1; -;  
DR PIR; A36421; MENZP2.  
DR PIR; S16661; S16661.  
InterPro; IPR000982; Matrix.  
DR Pfam; PF00661; Matrix.1.  
DR ProDom; PD000741; Matrix.1.  
KM Matrix protein; Envelope protein.  
SQ SEQUENCE 377 AA; 42312 MW; A181A79C55276736 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 377;  
Best Local Similarity 60.0%; Pred. No. 8.9;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSAVALTYSP 10  
|:|:|:|:|  
Db 170 PTFVSTYSP 179

RESULT 6  
ID POLN\_RHDV STANDARD; PRT; 2344 AA.  
AC P27410;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Non-structural polypeptide (contains: RNA-directed RNA polymerase  
protein; Coat protein).  
DE protein; Coat protein).  
OS Rabbit hemorrhagic disease virus (RHDV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Lagovirus.  
OX NCBI\_TaxID=11976;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91361557; PubMed=1840711;  
RA Meyers G., Wirblich C., Thiel H.-J.;  
RT "Rabbit hemorrhagic disease virus - molecular cloning and nucleotide  
sequencing of a calicivirus genome.";  
Virology 184:664-676 (1991).  
RL -1- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE  
FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.  
CC

```
CC -! CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -! SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M67473; AAA47285.1; -
DR PIR; A41039; RRMWRH.
DR MEROPS; C24.001; -.
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR004004; Calici pol hel.
DR InterPro; IPR000317; Endoptase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep RNA pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR Pfam; PF03510; Calici coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR PolyProtein; Transferrase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN 2508 2707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 21497 21625 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2144 COAT PROTEIN.
FT NP_BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 2344;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSAVALTYSP 10
|:|:|:|:|
Db 2194 PNASAITVTP 2203

RESULT 7
ADH7_HUMAN
ID ADH7_HUMAN STANDARD; PRT; 374 AA.
AC P40394; Q13713;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alcohol dehydrogenase class IV mu/sigma chain (EC 1.1.1.1) (Retinol
DE dehydrogenase) (Gastric alcohol dehydrogenase).
GN ADH7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RX MEDLINE=94253145; PubMed=8195208;
RA Satre M.A., Zgombic-Knight M., Duester G.;
RT "The complete structure of human class IV alcohol dehydrogenase
RT (retinol dehydrogenase) determined from the ADH7 gene.";
RL J. Biol. Chem. 269:15606-15612(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181413; PubMed=7876191;
RA Zgombic-Knight M., Foglio M.H., Duester G.;
RT "Genomic structure and expression of the ADH7 gene encoding human
```

```
RT class IV alcohol dehydrogenase, the form most efficient for retinol
RT metabolism in vitro.";
RL J. Biol. Chem. 270:4305-4311(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RX MEDLINE=95010033; PubMed=7925371;
RA Farres J., Moreno A., Crossas B., Peralba J.M., Allali Hassani A.,
RA Hjeltnest L., Joernvall H., Pares X.;
RT "Alcohol dehydrogenase of class IV (sigma sigma-ADH) from human
RT stomach. cDNA sequence and structure/function relationships.";
RL Eur. J. Biochem. 224:549-557(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RX MEDLINE=94364515; PubMed=8082805;
RA Yokoyama S., Matsuo Y., Rambotham R., Yokoyama R.;
RT "Molecular characterization of a class IV human alcohol dehydrogenase
RT gene (ADH7).";
RL FEBS Lett. 351:411-415(1994).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RX MEDLINE=95181314; PubMed=7876099;
RA Kedishvili N.Y., Boston W.F., Stone C.L., Hurley T.D., Peggs C.F.,
RA Thomasson H.R., Popov K.M., Carr L.G., Edenberg H.J., Li T.K.;
RT "Expression and kinetic characterization of recombinant human stomach
RT alcohol dehydrogenase. Active-site amino acid sequence explains
RT substrate specificity compared with liver isozymes.";
RL J. Biol. Chem. 270:3625-3630(1995).
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Fore skin;
RA Yokoyama H., Baraona E., Lieber C.S.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 227-299 FROM N.A.
RC TISSUE=Stomach;
RX MEDLINE=95289593; PubMed=7771649;
RA Cheung B., Anderson J.K., Holmes R.S., Beacham I.R.;
RT "Human stomach class IV alcohol dehydrogenase: molecular genetic
RT analysis.";
RL Alcohol. Clin. Exp. Res. 19:185-186(1995).
RN [8]
RP SEQUENCE OF 12-21; 41-46; 62-80; 85-111; 135-142 AND 169-174.
RX MEDLINE=92275085; PubMed=1592118;
RA Pares X., Cederlund E., Moreno A., Saubi N., Hooeog J.-O.,
RA Joernvall H.;
RT "Class IV alcohol dehydrogenase (the gastric enzyme). Structural
RT analysis of human sigma sigma-ADH reveals class IV to be variable and
RT confirms the presence of a fifth mammalian alcohol dehydrogenase
RT class.";
RL FEBS Lett. 303:69-72(1992).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=97373546; PubMed=9228021;
RA Xie P., Parsons S.H., Speckhard D.C., Bosron W.F., Hurley T.D.;
RT "X-ray structure of human class IV sigma sigma alcohol dehydrogenase.
RT Structural basis for substrate specificity.";
RL J. Biol. Chem. 272:18558-18563(1997)
CC -! FUNCTION: COULD FUNCTION IN RETINOL OXIDATION FOR THE SYNTHESIS OF
CC RETINOIC ACID, A HORMONE IMPORTANT FOR CELLULAR DIFFERENTIATION.
CC MEDIUM-CHAIN (OCTANOL) & AROMATIC (M-NITROBENZALDEHYDE) COMPOUNDS
CC ARE THE BEST SUBSTRATES. ETHANOL IS NOT A GOOD SUBSTRATE BUT AT
CC THE HIGH ETHANOL CONCENTRATIONS REACHED IN THE DIGESTIVE TRACT, IT
CC PLAYS A ROLE IN THE ETHANOL OXIDATION AND CONTRIBUTES TO THE FIRST
CC PASS ETHANOL METABOLISM.
CC -! CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
CC NADH.
CC -! COFACTOR: ZINC.
CC -! SUBUNIT: HOMODIMER.
CC -! SUBCELLULAR LOCATION: Cytoplasmic.
CC -! TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN STOMACH.
```

```

CC -1- MISCELLANEOUS: THERE ARE 7 DIFFERENT ADH'S ISOZYMES IN HUMAN:
CC THREE BELONGS TO CLASS-I: ALPHA, BETA, AND GAMMA, ONE TO CLASS-II:
CC PI, ONE TO CLASS-III: CH, ONE TO CLASS-IV: ADH7 AND ONE TO
CC CLASS-V: ADH6.
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY. CLASS-IV SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U16293; AAC51351.1; -
CC EMBL: U16286; AAC51351.1; JOINED.
CC EMBL: U16287; AAC51351.1; JOINED.
CC EMBL: U16288; AAC51351.1; JOINED.
CC EMBL: U16289; AAC51351.1; JOINED.
CC EMBL: U16290; AAC51351.1; JOINED.
CC EMBL: U16291; AAC51351.1; JOINED.
CC EMBL: U16292; AAC51351.1; JOINED.
CC EMBL: X76342; CAAS3960.1; -
CC EMBL: X76342; CAAS3960.1; ALT_INIT.
CC EMBL: U07821; AAA19002.1; -
CC EMBL: U09623; AAA82165.1; -
CC EMBL: L47166; AAB38424.1; -
CC EMBL: S77168; AAB34478.1; -
CC PIR: S21170; S21170.
CC PDB: 1AGN; 12-MAR-97.
CC Genew: HGNC:256; ADH7.
CC MIN: 600086; -
CC InterPro: IPR002328; ADH_zinc.
CC InterPro: IPR002085; Adh_zn_family.
CC Pfam: PF00107; adh_zinc; 1.
CC PROSITE: PS00059; ADH_ZINC; 1.
CC Oxidoreductase; Zinc; NAD; Multigene family; 3D-structure.
CC EMBL: 47; 47; ZINC (CATALYTIC) (BY SIMILARITY).
CC EMBL: 68; 68; ZINC (CATALYTIC) (BY SIMILARITY).
CC EMBL: 98; 98; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 101; 101; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 104; 104; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 112; 112; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 174; 174; ZINC (CATALYTIC) (BY SIMILARITY).
CC EMBL: 20; 20; Q -> R (IN REF. 5).
CC EMBL: 71; 71; T -> R (IN REF. 5).
CC EMBL: 98; 98; C -> V (IN REF. 6).
CC EMBL: 138; 139; HH -> GR (IN REF. 7).
CC EMBL: 148; 149; EX -> VT (IN REF. 6).
CC EMBL: 157; 157; V -> E (IN REF. 6).
CC EMBL: 205; 205; L -> V (IN REF. 6).
CC EMBL: 374; 374; AA; 40005 MW; 1D785CD022C8D387 CRC64;
CC SEQUENCE

```

```

Query Match 70.0%; Score 35; DB 1; Length 374;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 PSAVALITYSP 10
DB 296 PSAKMLTYDP 305

```

```

RESULT 8
ADH7_MOUSE STANDARD; PRT; 374 AA.
AC 064437;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alcoholic dehydrogenase class IV (EC 1.1.1.1) (retinol dehydrogenase)
DE (Gastric alcohol dehydrogenase) (ADH-C2) (Alcohol dehydrogenase 7).

```

```

GN ADH7 OR ADH3 OR ADH-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Gastric mucosa;
RX MEDLINE=95256259; Pubmed=7738026;
RA Zgonibic-Knight M., Ang H.L., Foglio M.H., Duester G.;
RT "Cloning of the mouse class IV alcohol dehydrogenase (retinol
RT dehydrogenase) cDNA and tissue-specific expression patterns of the
RT murine ADH gene family.";
RL J. Biol. Chem. 270:10868-10877(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ.
RX MEDLINE=97271563; Pubmed=9126489;
RA Zgonibic-Knight M., Deltour L., Haselbeck R.J., Foglio M.H.,
RA Duester G.;
RT "Gene structure and promoter for Adh-3 encoding mouse class IV alcohol
RT dehydrogenase (retinol dehydrogenase).";
RL Genomics 41:105-109(1997).
CC -1- FUNCTION: COULD FUNCTION IN RETINOL OXIDATION FOR THE SYNTHESIS OF
CC RETINOIC ACID, A HORMONE IMPORTANT FOR CELLULAR DIFFERENTIATION.
CC -1- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
CC NADH.
CC -1- COFACTOR: ZINC.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN THE STOMACH MUCOSA, LOWER
CC EXPRESSION IN EYE, THYMUS, SKIN AND OVARY. VERY LOW EXPRESSION IN
CC SMALL INTESTINE, LIVER AND UTERUS.
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY. CLASS-IV SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U20257; AAA76735.1; -
CC EMBL: U76734; AAC53124.1; JOINED.
CC EMBL: U76728; AAC53124.1; JOINED.
CC EMBL: U76729; AAC53124.1; JOINED.
CC EMBL: U76730; AAC53124.1; JOINED.
CC EMBL: U76727; AAC53124.1; JOINED.
CC EMBL: U76731; AAC53124.1; JOINED.
CC EMBL: U76733; AAC53124.1; JOINED.
CC EMBL: U76732; AAC53124.1; JOINED.
CC HSSP: P40394; 1AGN.
CC MGD: MGI:87926; Adh3.
CC InterPro: IPR002328; ADH_zinc.
CC InterPro: IPR002085; Adh_zn_family.
CC Pfam: PF00107; adh_zinc; 1.
CC PROSITE: PS00059; ADH_ZINC; 1.
CC Oxidoreductase; Zinc; NAD; Multigene family.
CC EMBL: 47; 47; ZINC (CATALYTIC) (BY SIMILARITY).
CC EMBL: 68; 68; ZINC (CATALYTIC) (BY SIMILARITY).
CC EMBL: 98; 98; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 101; 101; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 104; 104; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 112; 112; ZINC (SECOND ATOM) (BY SIMILARITY).
CC EMBL: 174; 174; ZINC (CATALYTIC) (BY SIMILARITY).
CC EMBL: 374; 374; AA; 39837 MW; 6E0978E44E02D027 CRC64;
CC SEQUENCE

```

```

Query Match 70.0%; Score 35; DB 1; Length 374;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```
QY 1 PSVAULTYSP 10
    |||
Db 296 PSAKMLTYDP 305

RESULT 9
RTG3 YEAST
ID RTG3 YEAST STANDARD; PRT; 486 AA.
AC P38165; P89494;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15 JUN-2002 (Rel. 41, Last annotation update)
DE Retrograde regulation protein 3.
GN RTG3 OR YBL101C OR YBL0810.
OS Saccharomyces cerevisiae [Baker's yeast].
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GRF88;
RX MEDLINE=97184436; PubMed=9032238;
WA Jia Y., Rothermel B., Thornton J., Butow R.A.;
RT "A basic helix-loop-helix-leucine zipper transcription complex in
RT yeast functions in a signaling pathway from mitochondria to the
RT nucleus."
RL Mol. Cell. Biol. 17:1110-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=96076635; PubMed=7502586;
RA Obermayer H., Gassenhuber J., Pitaravandi E., Domdey H.;
RT "Sequence analysis of a 78.6 kb segment of the left end of
RT Saccharomyces cerevisiae chromosome II."
RL Yeast 11:1103-1112(1995).
CC !- FUNCTION: TRANSCRIPTION FACTOR THAT REGULATES CIT2 GENE
CC EXPRESSION. BINDS TO TWO IDENTICAL SITES ORIENTED AS INVERTED
CC REPEATS 28 BP APART IN A REGULATORY UPSTREAM ACTIVATION SEQUENCE
CC ELEMENT (UASR) IN THE CIT2 PROMOTER. THE CORE BINDING SITE IS 5'-
CC GGTCAAC-3'.
CC ! SUBUNIT: BINDS TO DNA AS A HETERODIMER WITH RTG1.
CC ! SUBCELLULAR LOCATION: Nuclear (Probable).
CC ! SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC ! CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN AND IS SHORTER
CC (416 AA) DUE TO A FRAMESHIFT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U46012; AAA86842.1;
DR EMBL; X79489; CAA55992.1; ALT_FRAME.
DR EMBL; Z45464; CAA84929.1;
DR EMBL; Z45465; CAA84931.1;
DR PTR; S45392; S45392.
DR HSSP; P22415; 1AN4.
DR TRANSFAC; T04579;
DR SGD; S0000199; RTG3.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH1; 1.
DR PROSITE; PS00488; HLH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA BIND 283 298
FT DOMAIN 299 345
FT HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT SEQUENCE 486 AA; 54139 MW; D8635FDFEC740C11 CRC64;
```

```
Query Match 70.0%; Score 35; DB 1; Length 486;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSVAULTYSP 10
    |||
Db 214 PASVSSTYSP 223

RESULT 10
KRUI DROME
ID KRUI DROME STANDARD; PRT; 845 AA.
AC P08155; Q9VWH5;
DT 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Krueppel homologous protein 1.
GN KR-H1 OR KR-H OR CG9167.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND ALTERNATIVE SPLICING.
RC STRAIN=Oregon R; TISSUE=Embryo, Larva, and Pupae;
RX MEDLINE=20237577; PubMed=10772791;
RA Pecasse F., Beck Y., Ruiz C., Richards G.;
RT "Krueppel-homolog, a stage-specific modulator of the prepupal ecdysone
RT response, is essential for Drosophila metamorphosis."
RL Dev. Biol. 221:53-67(2000).
RN [2]
RP SEQUENCE FROM N.A. (ALPHA AND BETA ISOPFORMS).
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Balaz R.G., Champe M., Pfeiffer H.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.K., Miklos G.L.G.,
RA Abrial J.F., Abmayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Feriz C., Ferreira S., Fleischmann W.,
RA Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jialai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasse P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Poilard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
```

[1]  
 RN SEQUENCE OF 289-367 FROM N.A.  
 RP MEDLINE=87051757; PubMed=3096579;  
 RX Schuh R., Alcher W., Gaul U., Cote S., Preiss A., Maier D.,  
 RA Seifert E., Naber W., Schroeder C., Kemler R., Jaekle H.,  
 RT "A conserved family of nuclear proteins containing structural  
 RT elements of the finger protein encoded by Kruppel, a Drosophila  
 RT segmentation gene.";  
 RL Cell 47:1025-1032(1986).  
 CC -1- FUNCTION: PLAYS A GENERAL ROLE IN THE HIERARCHIES OF GENE  
 CC EXPRESSION LEADING TO METAMORPHOSIS.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOMORPHS; ALPHA AND BETA (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- DEVELOPMENTAL STAGE: BETA ISOFORM IS EXPRESSED DURING  
 CC EMBRYOGENESIS, MOST ABUNDANT IN MIDEMBRYOGENESIS, AND IN ADULTS.  
 CC ALPHA ISOFORM IS EXPRESSED FROM EMBRYOGENESIS TO 8 HOURS AFTER  
 CC PUPARIATION, MAJOR PERIOD OF EXPRESSION IS DURING SECOND INSTAR.  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ005440; CA006543.2; -  
 DR EMBL: AJ005441; CA006544.2; -  
 DR EMBL: AE003612; AF52343.2; -  
 DR EMBL: AE003612; AG22417.1; -  
 DR EMBL: M1940; AAA28660.1; -  
 DR HSSP: P08047; ISP2.  
 DR FlyBase: FBgn0028420; Kr-h1.  
 DR InterPro: IPR000822; Znf\_C2H2.  
 DR Pfam: PF00096; Znf\_C2H2\_8.  
 DR ProDom: PD000003; Znf\_C2H2\_4.  
 DR SMART: SM00355; Znf\_C2H2\_8.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 8.  
 DR PROSITE: PS00157; ZINC\_FINGER\_C2H2\_2; 8.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Alternative splicing;  
 KW Developmental protein.  
 FT DOMAIN 108 148 GLN-RICH.  
 FT DOMAIN 482 515 SER/THR-RICH.  
 FT DOMAIN 538 565 SER/THR-RICH.  
 FT DOMAIN 194 463 ZINC-FINGERS.  
 FT DOMAIN 194 463 ZINC-FINGERS.  
 FT ZN\_FING 271 293 C2H2-TYPE.  
 FT ZN\_FING 299 321 C2H2-TYPE.  
 FT ZN\_FING 327 349 C2H2-TYPE.  
 FT ZN\_FING 355 377 C2H2-TYPE.  
 FT ZN\_FING 383 407 C2H2-TYPE.  
 FT ZN\_FING 413 435 C2H2-TYPE.  
 FT ZN\_FING 441 463 C2H2-TYPE.  
 FT VARSPIC 1 54 MISSING (IN ISOFORM ALPHA).  
 FT CONFLICT 313 313 N -> D (IN REF. 3).  
 SO SEQUENCE 845 AA; 91451 MW; A4D87839DCB372 CRC64;

Query Match 70.0%; Score 35; DB 1; Length 845;  
 Best Local Similarity 60.0%; Pred. No. 34;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSAVALTTPS 10  
 DB 544 PSSUSTTSP 553

RESULT 11  
 CHS2\_YEAST STANDARD; PRT; 963 AA.  
 AC P14180;  
 DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl  
 DE transferase 2).  
 GN CHS2 OR YBR038W OR YBR0407.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=90143137; PubMed=2533436;  
 RX Silverman S.J.;  
 RT "Similar and different domains of chitin synthases 1 and 2 of S.  
 RT cerevisiae: two isozymes with distinct functions.";  
 RL Yeast 5:459-467(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=8288c;  
 RC Andre B., Cziepluch C., Hein C., Jauniaux J.C., Urrestarazu A.,  
 RA Vissers S.;  
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP MUTAGENESIS.  
 RX MEDLINE=95293998; PubMed=7775457;  
 RA Nagahashi S., Sudoh M., Ono N., Sawada R., Yamaguchi E., Uchida Y.,  
 RA Mio T., Takagi M., Arita M., Yamada-Okabe H.;  
 RT "Characterization of chitin synthase 2 of Saccharomyces cerevisiae.  
 RT Implication of two highly conserved domains as possible catalytic  
 RT sites.";  
 RL J. Biol. Chem. 270:13961-13967(1995).  
 CC -1- FUNCTION: ESSENTIAL FOR SEPTUM FORMATION AND CELL DIVISION. CHS2  
 CC IS REQUIRED FOR MAINTAINING NORMAL CELL MORPHOLOGY.  
 CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4) (N-acetyl)-  
 CC beta-D-glucosaminyl} (N) = UDP + {(1,4) (N-acetyl)-beta-D-  
 CC glucosaminyl} (N+1).  
 CC -1- ENZYME REGULATION: REQUIRES PROTEOLYTIC ACTIVATION.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M23865; AAA4493.1; -  
 DR EMBL: Z35907; CA484980.1; -  
 DR PIR: S45167; S45167.  
 DR PIR: A30922; A30922.  
 DR SGD: S0000242; CHS2.  
 DR InterPro: IPR004834; Chitin synth.  
 DR InterPro: IPR001173; Glycosyltransf\_2.  
 DR Pfam: PF01644; Chitin synth. 1.  
 DR ProDom: PD002998; Chitin synth. 1.  
 KW transferase; Glycosyltransferase; Transmembrane; Cell wall;  
 KW Multigene family.  
 FT DOMAIN 1 422 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 423 443 1 (POTENTIAL).  
 FT DOMAIN 444 643 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 644 664 2 (POTENTIAL).  
 FT DOMAIN 665 677 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 678 698 3 (POTENTIAL).  
 FT DOMAIN 699 711 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 712 732 4 (POTENTIAL).  
 FT DOMAIN 733 743 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 744 764 5 (POTENTIAL).  
 FT DOMAIN 765 775 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 776 796 6 (POTENTIAL).  
 FT DOMAIN 797 875 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 876 896 7 (POTENTIAL).





ANPC\_HUMAN STANDARD; PRT; 540 AA.  
 ID ANPC\_HUMAN STANDARD; PRT; 540 AA.  
 AC P17342;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Atrial natriuretic peptide clearance receptor precursor (ANP-C)  
 DE (ANP-C) (Atrial natriuretic peptide C-type receptor).  
 GN NRP3 OR ANP-C.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=90287735; PubMed=2162522;  
 RA Lowe D.G., Camerato T.R., Goeddel D.V.;  
 RT "cDNA sequence of the human atrial natriuretic peptide clearance  
 receptor.";  
 RL Nucleic Acids Res. 18:3412-3412(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=90386656; PubMed=2169733;  
 RA Porter D.G., Atkinson A., Fuller F., Miller J.A., Gregory L.C.,  
 Lewicki J.A.;  
 RT "Isolation and functional expression of the human atrial natriuretic  
 peptide clearance receptor cDNA.";  
 RL Biochem. Biophys. Res. Commun. 171:796-803(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lens epithelium;  
 RA Rae J.L., Shepard A.R.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases  
 CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. DOES NOT HAVE  
 GUANYLATE CYCLASE ACTIVITY.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.  
 CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO  
 WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)  
 WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE  
 CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
 CC -1- SIMILARITY: TO ANP-A AND ANP-B RECEPTORS IN THEIR EXTRACELLULAR  
 AND TRANSMEMBRANE DOMAINS.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: X52282; CA36523.1; -  
 EMBL: M59305; AAA51734.1; -  
 EMBL: AF025998; AAB88801.1; -  
 DR PIR: S10150; S10150.  
 DR PIR: A35896; A35896.  
 DR Genew: HGNC:7945; NRP3.  
 DR MIM: 108962; -  
 DR InterPro: IPR001828; ANP\_receptor.  
 DR InterPro: IPR001170; Ntpep\_receptor.  
 DR Pfam: PF01094; ANP\_receptor.1.  
 DR PRINTS: PR00255; NATPEPTIDR.  
 DR PROSITE: PS00458; ANP\_RECEPTORS.1.  
 DR Receptor: Glycoprotein; Transmembrane; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 540 ATRIAL NATRIURETIC PEPTIDE CLEARANCE  
 FT RECEPTOR.  
 FT DOMAIN 19 480 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 481 503 POTENTIAL.  
 FT DOMAIN 504 540 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 108 136 BY SIMILARITY.

FT DISULFID 213 261 BY SIMILARITY.  
 FT DISULFID 473 473 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 476 476 C -> SG (IN REF. 2).  
 SQ SEQUENCE 540 AA; 59766 MW; 53EE0020A296D6F5 CRC64;  
 Query Match  
 Best Local Similarity 68.0%; Score 34; DB 1; Length 540;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 PSAVALTSP 10  
 Db 2 PSLVLTSP 11  
 RESULT 14  
 ID DVL2\_MOUSE STANDARD; PRT; 736 AA.  
 AC Q60838;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Segment polarity protein dishevelled homolog DVL-2 (Dishevelled-2)  
 DE (DSH homolog 2).  
 GN DVL2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C; TISSUE=Brain;  
 RX MEDLINE=97042042; PubMed=8987313;  
 RA Klugensmith J., Yang Y., Axelrod J.D., Beier D.R., Ferrimon N.,  
 Sussman D.J.;  
 RT "Conservation of dishevelled structure and function between flies and  
 mice: isolation and characterization of Dvl2.";  
 RL Mech. Dev. 58:15-26(1996).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE SIGNAL TRANSDUCTION PATHWAY  
 MEDIATED BY MULTIPLE WNT GENES.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.  
 CC -1- SIMILARITY: BELONGS TO THE DSH FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: U24160; ARC52827.1; -  
 DR HSPF: Q12923; 3PDZ.  
 DR MGD: MGI:106613; DVL2.  
 DR InterPro: IPR000591; DEP.  
 DR InterPro: IPR001158; DIX.  
 DR InterPro: IPR003351; Dishevelled.  
 DR InterPro: IPR001478; PDZ.  
 DR Pfam: PF00595; PDZ.1.  
 DR Pfam: PF00610; DEP.1.  
 DR Pfam: PF00778; DIX.1.  
 DR Pfam: PF02377; Dishevelled.1.  
 DR ProDom: PD003639; DIX.1.  
 DR SMART: SM00021; DAX.1.  
 DR SMART: SM00049; DEP.1.  
 DR SMART: SM00228; PDZ.1.  
 DR PROSITE: PS50186; DEP.1.  
 DR PROSITE: PS50106; PDZ.1.

KW Developmental protein.  
 FT DOMAIN 7 12 POLY-GLY.  
 FT DOMAIN 112 122 POLY-PRO.  
 FT DOMAIN 235 240 POLY-ARG.  
 FT DOMAIN 267 339 PDZ.  
 FT DOMAIN 433 507 DEP.  
 FT DOMAIN 686 691 POLY-PRO.  
 SQ SEQUENCE 736 AA; 78802 MW; 9237373D3D3C92A60 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 736;  
 Best Local Similarity 60.0%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSVAULTYSP 10  
 Db 670 PSEMAPYNP 679

RESULT 15  
 ID BL14 CAEEL STANDARD; PRT; 943 AA.  
 AC P51559; 044762; 044763; 044764; 044765; 044766;  
 DT 01-OCT-1996 (rel. 34, Created)  
 DT 15-JUN-2002 (rel. 41, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE Endoprotease bli-4 precursor (EC 3.4.21.-) (Blisterase) (Blistered  
 DE cuticle protein 4).  
 GN BLI-4 (OR K04F10.4).  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D), AND FUNCTION.  
 RC STRAIN Bl:stol N2;  
 RX MEDLINE:96293228; PubMed=7774813;  
 RA Thacker C., Peters K.W., Srayko M., Rose A.M.;  
 RT "The bli 4 locus of Caenorhabditis elegans encodes structurally  
 RT distinct kex2/subtilisin-like endoproteases essential for early  
 RT development and adult morphology";  
 RL Genes Dev. 9:956-971(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN:Bl:stol N2;  
 RA Latreille P., Wansley P.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 CC ! FUNCTION: Isoform A is required for normal production or  
 CC maintenance of the adult cuticle. There is functional redundancy  
 CC between the isoforms. Vital for embryonic and larval development.  
 CC !- TISSUE SPECIFICITY: In larvae and adults, expressed in all  
 CC hypodermal cells, vulva and ventral nerve cords.  
 CC !- DEVELOPMENTAL STAGE: Expression starts at embryo two-fold stage  
 CC through to adults.  
 CC !- ALTERNATIVE PRODUCTS: 5 isoforms; A/a, B/b, C/c, D/d (shown here)  
 CC and e, may be produced by alternative splicing.  
 CC !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. PURIN SUBFAMILY.  
 CC !- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 651.  
 CC !- CAUTION: Ref.2 (AAB96754/BA96757) sequence differs from that  
 CC shown due to erroneous gene model prediction.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; L29438; AAA98750.1; ALT\_FRAME.  
 EMBL; L29439; AAA98751.1; ALT\_FRAME.  
 EMBL; L29440; AAA98752.1; ALT\_FRAME.  
 EMBL; AF039719; AAB96753.1; -

DR EMBL; AF039719; AAB96754.1; ALT\_SEQ.  
 DR EMBL; AF039719; AAB96755.1; -  
 DR EMBL; AF039719; AAB96756.1; -  
 DR EMBL; AF039719; AAB96757.1; ALT\_SEQ.  
 DR MEROPS; S08.031; -  
 DR WormPep; K04F10.4a; CE11728.  
 DR WormPep; K04F10.4b; CE11730.  
 DR WormPep; K04F10.4c; CE11732.  
 DR WormPep; K04F10.4d; CE11734.  
 DR WormPep; K04F10.4e; CE11736.  
 DR InterPro; IPR002884; P domain.  
 DR InterPro; IPR002029; Peptidase\_S8.  
 DR Pfam; PF00082; Peptidase\_S8; 1.  
 DR Pfam; PF01483; P; PARTIAL.  
 DR PRINTS; PR00723; SUBTILISIN.  
 DR PROSITE; PS00136; SUBTILASE ASP; 1.  
 DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
 DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
 DR Hydrolase; Serine protease; Glycoprotein; Calcium-binding; Zymogen;  
 KW Transmembrane; Signal; Alternative splicing.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT PROPEP 21 116 BY SIMILARITY.  
 FT CHAIN 117 943 ENDOPROTEASE BLI-4.  
 FT ACT\_SITE 202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 241 241 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 415 415 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT TRANSMEM 417 437 POTENTIAL.  
 FT DOMAIN 438 684 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 713 857 CYS-RICH.  
 FT DOMBOHYD 195 195 N-LINKED (GLUCNA... (POTENTIAL).  
 FT VARSPPLIC 658 670 VEESARSPFDLT -> ILITIAHLVNA (IN  
 FT VARSPPLIC 658 699 ISOFORM A).  
 FT VARSPPLIC 658 699 VEESARSPFDLTSGMKLSQDCNGCTESSATSCFAYKH  
 FT L -> LCFNFENSQTECVLCVCHVPGVINFQIEKLLKST  
 FT KCLIFS (IN ISOFORM E).  
 FT VEESARSPFDLTSGMKLSQDCNGCTESSATSCFAYKH  
 FT LTQTLNKGSGFKVQKQCDTYYLDGKCKM -> GDEVV  
 FT ERIRHWEVTLESSHWNHAREHKSLOHLNSSKRTISPL  
 FT YSFTKFPQIFLILVCFIDAIHQFAV (IN ISOFORM  
 FT B).  
 FT VARSPPLIC 671 943 MISSING (IN ISOFORM A).  
 FT VARSPPLIC 700 943 MISSING (IN ISOFORM E).  
 FT VARSPPLIC 731 943 MISSING (IN ISOFORM B).  
 FT VARSPPLIC 780 828 ESNLVQAKIMRKDLGCGYIINAVGKLCUSSFTCTAP  
 FT GPMSCCKC -> AENPFCANNESGRTIVFKFKKRAAF  
 FT KDCYNLKLVDLDFLFLSLF (IN ISOFORM C).  
 FT VARSPPLIC 829 943 MISSING (IN ISOFORM C).  
 FT CONFLICT 153 153 A -> R (IN REF. 2).  
 FT CONFLICT 346 346 S -> R (IN REF. 2).  
 SQ SEQUENCE 943 AA; 103146 MW; 401E009E6C46AD7E CRC64;

Query Match 68.0%; Score 34; DB 1; Length 943;  
 Best Local Similarity 66.7%; Pred. No. 62;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSVAULTYSP 9  
 Db 381 PSSIAITYS 389

RESULT 16  
 ID CINS\_HUMAN STANDARD; PRT; 2016 AA.  
 AC Q14524;  
 DT 15-DEC-1998 (rel. 37, Created)  
 DT 15-DEC-1998 (rel. 37, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Sodium channel protein, cardiac muscle alpha-subunit (HHL1).  
 GN SCN5A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=92115699; PubMed=1309946;  
 RA Gellens M.E., George A.L. Jr., Chen L.Q., Chahine M., Horn R.,  
 RT Barclay R.L., Kallen R.G.;  
 "Primary structure and functional expression of the human cardiac  
 L-type calcium channel  $\alpha_1$  subunit.";  
 Proc. Natl. Acad. Sci. U.S.A. 89:554-558(1992).  
 RN [2]  
 RP VARIANTS LQTS.  
 RX MEDLINE=95196273; PubMed=7889574;  
 RA Wang Q., Shen J., Splawski I., Atkinson D., Li Z., Robinson J.L.,  
 RT Moss A.J., Towbin J.A., Keating M.T.;  
 "SCN5A mutations associated with an inherited cardiac arrhythmia, long  
 QT syndrome.";  
 Cell 80:805-811(1995).  
 RN [3]  
 RP VARIANTS LQTS.  
 RX MEDLINE=96081224; PubMed=8541846;  
 RA Wang Q., Shen J., Li Z., Timothy K., Vincent G.M., Priori S.G.,  
 RT Schwartz P.J., Keating M.T.;  
 "Cardiac sodium channel mutations in patients with long QT syndrome,  
 an inherited cardiac arrhythmia.";  
 Hum. Mol. Genet. 4:1603-1607(1995).  
 RN [4]  
 RP VARIANTS LQTS 1505-LYS--GLN-1507 DEL.  
 RX MEDLINE=95179949; PubMed=7651517;  
 RA Bennett P.B., Yazawa K., Makita N., George A.L. Jr.;  
 RT "Molecular mechanism for an inherited cardiac arrhythmia.";  
 Nature 376:683-685(1995).  
 RN [5]  
 RP VARIANTS LQTS GLY-1790.  
 RX MEDLINE=98349542; PubMed=9686753;  
 RA An R.H., Wang X.L., Kelem B., Benhorin J., Medina A., Goldmit M.,  
 RT Koss R.S.;  
 "Novel LQT-3 mutation affects Na<sup>+</sup> channel activity through  
 interactions between alpha- and beta-subunits.";  
 Circ. Res. 83:141-146(1998).  
 RN [6]  
 RP VARIANTS LQTS GLN-1623.  
 RX MEDLINE=9815676; PubMed=9506831;  
 RA Makita N., Shirai N., Nagashima M., Matsuo K., Yamada Y., Tohse N.,  
 RT Kitabatake A.;  
 "A de novo missense mutation of human cardiac Na<sup>+</sup> channel exhibiting  
 novel molecular mechanisms of long QT syndrome.";  
 FEBS Lett. 423:5-9(1998).  
 RN [7]  
 RP VARIANTS LQTS GLY-1839.  
 RA Benhorin J., Goldmit M., Maccluer J.M., Blangero J., Goffen R.,  
 RT Leibovitch A., Rahat A., Wang Q., Medina A., Towbin J., Kelem B.;  
 "Identification of a new SCN5A mutation, D1840G, associated with the  
 long QT syndrome.";  
 Hum. Mutat. 12:72-72(1998).  
 RN [8]  
 RP VARIANTS LQTS GLN-1623.  
 RA Yamagishi H., Furutani M., Kamisago M., Morikawa Y., Kojima Y.,  
 RT Hino Y., Furutani Y., Kimura M., Imamura S.-I., Takao A., Momma K.,  
 RX Matsuo K.;  
 "A de novo missense mutation (R1623Q) of the SCN5A gene in a Japanese  
 girl with sporadic long QT syndrome.";  
 Hum. Mutat. 12:481-481(1998).  
 RN [9]  
 RP VARIANTS LQTS MET-1304 AND MET-1645, AND VARIANT ASN-1500.  
 RX MEDLINE=99439526; PubMed=10508990;  
 RA Watanabe-Ichikawa D., Vesely M.R., Duggal P., Levine J.C.,  
 RT Blume E.D., Wolff G.S., Edwards S.B., Beggs A.H.;  
 "Sodium channel abnormalities are frequent in patients with long QT  
 syndrome: identification of two novel SCN5A mutations.";  
 Am. J. Med. Genet. 86:470-476(1999).  
 RN [10]  
 RP VARIANTS LQTS LYS-1784.

RX MEDLINE=99307063; PubMed=10377081;  
 RA Wei J., Wang D.W., Alings M., Fish F., Wathen M., Roden D.M.,  
 RA George A.L. Jr.;  
 RT "Congenital long-QT syndrome caused by a novel mutation in a conserved  
 acidic domain of the cardiac Na<sup>+</sup> channel.";  
 Circulation 99:3165-3171(1999).  
 RN [11]  
 RP VARIANTS LQTS N-1114; V-1501; L-1623; H-1644 AND N-1787.  
 RX MEDLINE=20432616; PubMed=10973849;  
 RA Splawski I., Shen J., Timothy K., Lehmann M.H., Priori S.,  
 RT Robinson J.L., Moss A.J., Schwartz P.J., Towbin J.A., Vincent G.M.,  
 RT Keating M.T.;  
 "Spectrum of mutations in long-QT syndrome genes. KVLQT1, HERG, SCN5A,  
 RT KCNE1, and KCNE2.";  
 Circulation 102:1178-1185(2000).  
 RN [12]  
 RP FUNCTION: THIS PROTEIN MEDIATES THE VOLTAGE-DEPENDENT SODIUM ION  
 PERMEABILITY OF EXCITABLE MEMBRANES, ASSUMING OPENED OR CLOSED  
 CONFORMATIONS IN RESPONSE TO THE VOLTAGE DIFFERENCE ACROSS THE  
 MEMBRANE. THE PROTEIN FORMS A SODIUM-SELECTIVE CHANNEL THROUGH  
 WHICH NA<sup>+</sup> IONS MAY PASS IN ACCORDANCE WITH THEIR ELECTROCHEMICAL  
 GRADIENT. IT IS A TETRODOTOXIN-RESISTANT NA<sup>+</sup> CHANNEL ISOFORM. THIS  
 CHANNEL IS RESPONSIBLE FOR THE INITIAL UPSTROKE OF THE ACTION  
 POTENTIAL IN THE ELECTROCARDIOGRAM.  
 CC SUBCELLULAR LOCATION: Integral membrane protein.  
 CC TISSUE SPECIFICITY: EXPRESSED IN HUMAN ATRIAL AND VENTRICULAR  
 CARDIAC MUSCLE BUT NOT IN ADULT SKELETAL MUSCLE, BRAIN,  
 MYOMETRIUM, LIVER, OR SPLEEN.  
 CC DOMAIN: THE SEQUENCE CONTAINS 4 INTERNAL REPEATS, EACH WITH 5  
 HYDROPHOBIC SEGMENTS (S1,S2,S3,S5,S6) AND ONE POSITIVELY CHARGED  
 SEGMENT (S4). SEGMENTS S4 ARE PROBABLY THE VOLTAGE-SENSORS AND ARE  
 CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT  
 EVERY THIRD POSITION.  
 CC DISEASE: DEFECTS IN SCN5A ARE THE CAUSE OF LONG QT SYNDROME TYPE 3  
 (LQTS). AN AUTOSOMAL DOMINANT CARDIAC DISEASE CHARACTERIZED BY  
 RECURRENT SYNCOPE AND SUDDEN CARDIAC DEATH.  
 CC MISCELLANEOUS: NA<sup>+</sup> CHANNELS IN MAMMALIAN CARDIAC MEMBRANE HAVE  
 FUNCTIONAL PROPERTIES QUITE DISTINCT FROM NA<sup>+</sup> CHANNELS IN NERVE  
 AND SKELETAL MUSCLE.  
 CC SIMILARITY: TO OTHER SODIUM CHANNEL PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
 CC -1- DATABASE: NAME=LQTSdb, NOTE=SCN5A mutations page:  
 WWW="http://www.ssi.dk/en/forkeking/lqtsdb/scn5a.htm".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL: M77235; AAA5644.1; -  
 DR Genbank: HGNC:10593; SCN5A.  
 DR MIM: 600163; -  
 DR InterPro: IPR001682; Ca/Na\_pore.  
 DR InterPro: IPR002111; Cat channel\_Tryp.  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR000636; M\_channel\_nlg.  
 DR Pfam: PF00520; Ion\_Tran\_4.  
 DR Pfam: PF00612; IQ\_1.  
 DR PRINTS: PRO0170; NACHANNEL.  
 DR SMART: SM00015; IQ\_1.  
 KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;  
 KW Glycoprotein; Repeat; Multigene family; Phosphorylation;  
 KW Polynorphism; Disease mutation; Long QT syndrome.  
 FT TRANSMEM 127 150  
 FT TRANSMEM 159 178  
 FT TRANSMEM 192 210  
 FT TRANSMEM 217 236  
 FT TRANSMEM 253 276  
 FT TRANSMEM 390 415  
 FT TRANSMEM 712 736  
 S1 OF REPEAT I (POTENTIAL).  
 S2 OF REPEAT I (POTENTIAL).  
 S3 OF REPEAT I (POTENTIAL).  
 S4 OF REPEAT I (POTENTIAL).  
 S5 OF REPEAT I (POTENTIAL).  
 S6 OF REPEAT I (POTENTIAL).  
 S1 OF REPEAT II (POTENTIAL).

```

FT TRANSMEM 748 771 S2 OF REPEAT I1 (POTENTIAL).
FT TRANSMEM 780 799 S3 OF REPEAT I1 (POTENTIAL).
FT TRANSMEM 806 825 S4 OF REPEAT I1 (POTENTIAL).
FT TRANSMEM 842 862 S5 OF REPEAT I1 (POTENTIAL).
FT TRANSMEM 914 939 S6 OF REPEAT I1 (POTENTIAL).
FT TRANSMEM 1201 1224 S1 OF REPEAT I11 (POTENTIAL).
FT TRANSMEM 1238 1263 S2 OF REPEAT I11 (POTENTIAL).
FT TRANSMEM 1270 1291 S3 OF REPEAT I11 (POTENTIAL).
FT TRANSMEM 1296 1317 S4 OF REPEAT I11 (POTENTIAL).
FT TRANSMEM 1337 1359 S5 OF REPEAT I11 (POTENTIAL).
FT TRANSMEM 1444 1470 S6 OF REPEAT I11 (POTENTIAL).
FT TRANSMEM 1524 1547 S1 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1559 1582 S2 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1589 1612 S3 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1623 1644 S4 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1660 1682 S5 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1748 1772 S6 OF REPEAT IV (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 548 548 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 592 592 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 740 740 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 803 803 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 841 841 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 864 864 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1365 1365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1374 1374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1380 1380 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1388 1388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1736 1736 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1774 1774 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1955 1955 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 558 558 H -> R.
FT VARIANT 1114 1114 /FTID=VAR_008955.
FT VARIANT 1304 1304 D -> N (IN LQT3).
FT VARIANT 1404 1404 /FTID=VAR_009935.
FT VARIANT 1425 1425 T -> M (IN LQT3).
FT VARIANT 1425 1425 /FTID=VAR_008956.
FT VARIANT 1425 1425 N -> S (IN LQT3).

Query Match 68.0%; Score 34; DB 1; Length 1616;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PSAAATYSP 10
||| : |||
Db 1002 PSCIATPSP 1011

RESULT 17
ID PYRO SULAC STANDARD; PRT; 291 AA.
AC 008958;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Dihydroorotate dehydrogenase (EC 1.3.3.1) (Dihydroorotate oxidase)
DE (DHODase) (DHOD).
CN PYRO.
OS Sulfolobus acidocaldarius.
OC Archaea; Euryarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
CX NCBI_TaxID=2285;
CN 11
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 31909 / NCBI 11770 / DSM 639;
RA Charlier D.R.M., Thia-Toong T.L., Roovers M., Durbecq V., Legrain C.,
RA Glandsdorff N.;

Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
!- CATALYTIC ACTIVITY: (S)-dihydroorotate + O(2) -> orotate +
H(2)O(2).
!- COFACTOR: FMN (BY SIMILARITY).
!- PATHWAY: Pyrimidine biosynthesis; fourth step.
!- SUBUNIT: HOMODIMER (BY SIMILARITY).
!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
!- SIMILARITY: BELONGS TO THE DIHYDROOROTATE DEHYDROGENASE FAMILY.
SUBFAMILY 1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Y12823; CAAY73354.1; ALT_INIT.
InterPro; IPR001295; DHO dh.
InterPro; IPR003009; FMN enzyme.
Pfam; PF01180; DHODase; 1.
TIGRfam; TIGR01037; PYRD subf. fam. 1.
PROSITE; PS00911; DHODASE_1; FALSE_NEG.
PROSITE; PS00912; DHODASE_2; 1.
Pyrimidine biosynthesis; Oxidoreductase; Flavoprotein; FMN.
FT NP BIND 224 232 FMN (POTENTIAL).
SQ SEQUENCE 291 AA; 31667 MW; 90C624D3BB0D30ED CRC64;

Query Match 67.0%; Score 33.5; DB 1; Length 291;
Best Local Similarity 61.5%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

Qy 1 PSAAATYSP 10
||| : |||
Db 36 PSAITKTLTSP 48

RESULT 18
ID GBLP SCHPO STANDARD; PRT; 314 AA.
AC Q108J1; P78896;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Guanine nucleotide-binding protein beta subunit-like protein (Receptor
of activated protein kinase C).
CN RKP1 OR CPC2 OR SPAC6B12.15.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
CX NCBI_TaxID=4896;
CN 11
RP SEQUENCE FROM N.A.
RA Park S.-K., Yoo H.-S.;
RA "Molecular cloning and nucleotide sequence of Schizosaccharomycetes
pombe homologue of the receptor for activated protein kinase C gene.";
Korean J. Microbiol. 33:128-131(1995).
[2]
RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
RC STRAIN=ED616;
RX MEDLINE=21164691; PubMed=11263963;
RA Won M., Park S.-K., Hoe K.-L., Jang Y.-J., Chung K.-S., Kim D.-U.,
RA Kim H.-B., Yoo H.-S.;
RT "Rkp1/Cpc2, a fission yeast RACK1 homolog, is involved in actin
cytoskeleton organization through protein kinase C, Pck2, signaling.";
RL Biochem. Biophys. Res. Commun. 282:10-15(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;

```

RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Spouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown S., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Wolckele G., Aert R., Robben J., Gympre B.,  
 RA Welyens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Manhart R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Ares S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,  
 RA Cerretti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovskiy G.V., Ussey D., Barrell B.G., Nurse P.,  
 RA "The genome sequence of Schizosaccharomyces pombe."  
 Nature 415:871-880(2002).  
 [4]  
 RA SEQUENCE OF 5-314 FROM N.A.  
 RP STRAIN=PR745;  
 RX MEDLINE=98162722; PubMed=9501991;  
 RT Yoshiooka S., Kato K., Nakai K., Okayama H., Nojima H.;  
 RT "Identification of open reading frames in Schizosaccharomyces pombe  
 CDNA's";  
 RL DNA Res. 4:363-369(1997).  
 CC -1- FUNCTION: MAY BE A RECEPTOR FOR PROTEIN KINASE C IN THE REGULATION  
 CC OF ACTIN CYTOSKELETON ORGANIZATION DURING CELL WALL SYNTHESIS AND  
 CC MORPHOGENESIS.  
 CC -1- SUBUNIT: INTERACTS WITH PKC2.  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH PARTICULATE FRACTIONS.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, L37885; AAAS6865.2; -;  
 DR EMBL, AF320333; AAK38633.1; -;  
 DR EMBL, Z98531; CAB11079.1; -;  
 DR EMBL, D89247; BAA13908.1; -;  
 DR InterPro, IPR001680; WD40.  
 DR Pfam, PF00400; WD40; 7.  
 DR PRINTS, PR00320; GPROTEINRPT.  
 DR ProDom, PD000018; WD40; 5.  
 DR SMART, SMO0320; WD40; 7.  
 DR PROSITE, PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE, PS00678; WD\_REPEATS\_2; 6.  
 DR PROSITE, PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD Repeat.  
 FT REPEAT 13 44 WD 1.  
 FT REPEAT 61 91 WD 2.  
 FT REPEAT 103 133 WD 3.  
 FT REPEAT 146 178 WD 4.  
 FT REPEAT 190 220 WD 5.  
 FT REPEAT 231 260 WD 6.  
 FT REPEAT 281 311 WD 7.  
 FT REPEAT 41 41 I -> L (IN REF. 1).  
 SO SEQUENCE 314 AA; 34851 MW; 4E14707164E68ACD CRC64;

Query Match

66.0%; Score 33; DB 1; Length 314;

Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 PSAVALTYSP 10  
 Db 284 PECISLTWSP 293  
 RESULT 19  
 VVAT LPV  
 ID VVAT LPV STANDARD; PRT; 369 AA.  
 AC P27287;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DE 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Matrix protein.  
 GN M.  
 OS La Piedad-Michoacan-Mexico virus (LPV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91237338; PubMed=2033388;  
 RA Berg M., Sundqvist A., Moreno-Lopez J., Linares T.;  
 RT "Identification of the porcine paramyxovirus LPV matrix protein  
 RT gene: comparative sequence analysis with other paramyxoviruses";  
 RL J. Gen. Virol. 72:1045-1050(1991).  
 CC -1- FUNCTION: THE M PROTEIN HAS A CRITICAL ROLE IN VIRUS ASSEMBLY  
 CC AND INTERACTS WITH THE RNP COMPLEX AS WELL AS WITH THE VIRAL  
 CC MEMBRANE.  
 CC PIR, A38707; MFZLA.  
 DR InterPro, IPR000982; Matrix.  
 DR Pfam, PF00661; Matrix; 1.  
 DR ProDom, PD000741; Matrix; 1.  
 KW Matrix protein; Envelope protein.  
 SQ SEQUENCE 369 AA; 41657 MW; B5362722FF8D6B5A CRC64;  
 Query Match 66.0%; Score 33; DB 1; Length 369;  
 Best Local Similarity 60.0%; Pred. No. 35;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 PSAVALTYSP 10  
 Db 173 PRFVSLTYCP 182  
 RESULT 20  
 ADHA UROHA  
 ID ADHA UROHA STANDARD; PRT; 375 AA.  
 AC P25405;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Alcohol dehydrogenase I-A (EC 1.1.1.1) (ADH IAI).  
 OS Uromastix hardwickii (Indian spiny-tailed lizard).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Iguania; Acrodontia; Agamidae; Uromastixinae;  
 OC Uromastix.  
 OX NCBI\_TaxID=40250;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96195666; PubMed=8612630;  
 RA Hjeltnyist L., Shafiqat J., Siddiqi A.R., Oerenvall H.;  
 RT "Linking of isozyme and class variability patterns in the emergence  
 RT of novel alcohol dehydrogenase functions. Characterization of  
 RT isozymes in Uromastix hardwickii";  
 RL Eur. J. Biochem. 235:563-570(1996).  
 RN [2]  
 RP SEQUENCE OF 1-18.  
 RC TISSUE=Liver;  
 RX MEDLINE=92183888; PubMed=1544464;  
 RA Hjeltnyist L., Ericsson M., Shafiqat J., Carlquist M., Siddiqi A.R.,

RA Hooey J. G., Joernvall H.;  
 RT "Reptilian alcohol dehydrogenase. Heterogeneity relevant to class  
 RL multiplicity of the mammalian enzyme.";  
 RL FEBS Lett. 298:297-300(1992).  
 CC ! CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +  
 CC NADH.  
 CC ! COFACTOR: ZINC.  
 CC ! SUBUNIT: MULTIMERIC (WITH DIFFERENT RATIOS OF MONOMERS).  
 CC ! SUBCELLULAR LOCATION: Cytoplasmic.  
 CC ! MISCELLANEOUS: IN U.HARDWICKII THERE ARE TWO ISOZYMES OF ALCOHOL  
 CC DEHYDROGENASE I.  
 CC ! SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE  
 CC FAMILY.  
 DR PIR: S20429; S20429.  
 DR HSSP: P06127; ZOHX.  
 DR InterPro: IPR002328; ADH\_zinc.  
 DR InterPro: IPR002085; Adh\_zn\_family.  
 DR Pfam: PF00107; adh\_zinc; 1.  
 DR PROSITE: PS00559; ADH\_ZINC; 1.  
 KW (oxidoreductase; Zinc; NAD; Multigene family; Acetylation.  
 FT MOD RES 1 1  
 FT METAL 46 46 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 67 67 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 97 97 ZINC (SECOND ATOM) (BY SIMILARITY).  
 FT METAL 100 100 ZINC (SECOND ATOM) (BY SIMILARITY).  
 FT METAL 103 103 ZINC (SECOND ATOM) (BY SIMILARITY).  
 FT METAL 111 111 ZINC (SECOND ATOM) (BY SIMILARITY).  
 FT METAL 174 174 ZINC (CATALYTIC) (BY SIMILARITY).  
 SQ SEQUENCE 475 AA; 39663 MW; F315BB777996D906 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 375;

Best Local Similarity 60.0%; Pred. No. 36;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSVAULTYSP 10  
 ||| ||| |||  
 Db 297 PSASHTFSP 306

RESULT 21

VMAT SV41  
 ID VMAT SV41 STANDARD; PRT: 382 AA.  
 AC P25182;  
 DT 01 MAY-1992 (Rel. 22, Created)  
 DT 01 MAY 1992 (Rel. 22, Last sequence update)  
 DT 15-DEC 1998 (Rel. 37, Last annotation update)  
 DE Matrix protein.  
 GN M.  
 OS Simian virus 41 (SV41).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.  
 CC NCBI\_TaxID=11228;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Toshiba/Chanock;  
 RX MEDLINE=91049440; PubMed=2173260;  
 RA Tsurudome M., Bando H., Nishio M., Iwamoto Y., Kawano M., Kondo K.,  
 RA Komada H., Ito Y.;  
 ET "Antigenic and structural properties of a paramyxovirus simian virus  
 RT 41 (SV41) reveal a close relationship with human parainfluenza type 2  
 RL virus.";  
 RL Virology 179:738-748(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Toshiba/Chanock;  
 RX MEDLINE=91135794; PubMed=1651608;  
 RA Tsurudome M., Bando H., Kawano M., Matsumura H., Komada H., Nishio M.,  
 RA Ito Y.;  
 ET "Transfer of simian virus 41 (SV41) matrix gene are exclusively  
 RT dicistronic with the fusion gene which is also transcribed as a  
 RT monocistron.";  
 RL Virology 184:94-100(1991).  
 CC ! FUNCTION: THE M PROTEIN HAS A CRUCIAL ROLE IN VIRUS ASSEMBLY

CC AND INTERACTS WITH THE RNP COMPLEX AS WELL AS WITH THE VIRAL  
 CC MEMBRANE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: S48627; AAB19493.1; -;  
 DR EMBL: X64275; CAA45566.1; -;  
 DR EMBL: M62733; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A40563; MFNZ41.  
 DR InterPro: IPR000982; Matrix.  
 DR Pfam: PF00661; Matrix; 1.  
 DR ProDom: PD000741; Matrix; 1.  
 KW Matrix protein; Envelope protein.

SQ SEQUENCE 382 AA; 42155 MW; B4C6DCE968E4A855 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 382;

Best Local Similarity 60.0%; Pred. No. 36;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSVAULTYSP 10  
 ||| ||| |||  
 Db 170 PTFVSLTYCP 179

RESULT 22

EFIA DESMO  
 ID EFIA DESMO STANDARD; PRT: 438 AA.  
 AC P41203;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Elongation factor 1-alpha (EF-1-alpha) (Elongation factor Tu) (EF Tu).  
 GN TUF OR TUX.  
 OS Desulfurococcus mobilis.  
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
 OC Desulfurococcaceae; Desulfurococcus.  
 CC NCBI\_TaxID=2274;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 2126;  
 RX MEDLINE=95206243; PubMed=7898436;  
 RA Ceccarelli E., Bocchetta M., Creti R., Sanangelantoni A.M., Tiboni O.,  
 RA Cammarano F.;  
 ET "Chromosomal organization and nucleotide sequence of the genes for  
 RT elongation factors EF-1 alpha and EF-2 and ribosomal proteins S7 and  
 RT S10 of the hyperthermophilic archaeum Desulfurococcus mobilis.";  
 RL Mol. Gen. Genet. 246:687-696(1995).  
 CC !- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
 CC BIOSYNTHESIS.  
 CC !- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC !- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC EF-TU/EF-1A SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X73582; CAA51984.1; -;  
 DR HSSP: P07157; IAIP.  
 DR InterPro: IPR004539; EFL alpha.  
 DR InterPro: IPR004160; EFTU Cterm.  
 DR InterPro: IPR004161; EFTU\_D2.

DR InterPro: IPR000795; EF\_GTPbind.  
 DR Pfam: PF00009; GTP\_EFTU; 1.  
 DR Pfam: PF03143; GTP\_EFTU\_D3; 1.  
 DR Pfam: PF03144; GTP\_EFTU\_D2; 1.  
 DR PRINTS: PRO0315; ELONGATN\_FCT.  
 DR TIGRPFAM: TIGR00483; EF-1\_alpha; 1.  
 DR PROSITE: PS00301; EFATOR\_GTP; 1.  
 DR Elongation factor; Protein biosynthesis; GTP-binding.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 92 96 GTP (BY SIMILARITY).  
 FT NP\_BIND 154 157 GTP (BY SIMILARITY).  
 SQ SEQUENCE 438 AA; 46737 MW; 4A4955EB6C39427 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 438;  
 Best Local Similarity 50.0%; Pred. No. 42;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSVAITYSP 10  
 DB 338 PTAIAVGYTP 347

RESULT 23  
 FIBI\_ADE40 STANDARD; PRT; 547 AA.

PI8047;  
 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE Fiber protein 1.  
 OS Human adenovirus type 40.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=28284;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=89370295; PubMed=2773314;  
 RA Kidd A.H.; Erasmus M.J.;  
 RT "Sequence characterization of the adenovirus 40 fiber gene."  
 RT Virology 173:134-144(1989).  
 CC - FUNCTION: RECOGNIZES THE CELL RECEPTOR; SERVES AS THE LIGAND  
 CC BETWEEN THE ADENOVIRUS CAPSID AND THE HOST CELL RECEPTOR.  
 CC - SUBUNIT: HOMOTRIMER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M28822; AAA03234.1; -;  
 DR EMBL: L19443; AAC13980.1; -;  
 DR PIR: B30336; ERADEF4.  
 DR HSSP: P11818; 1XNB.  
 DR InterPro: IPR000939; Adeno\_fiber2.  
 DR InterPro: IPR000978; Adeno\_fiber\_knob.  
 DR InterPro: IPR000931; Adeno\_fiber.  
 DR Pfam: PF00541; adeno\_fiber; 1.  
 DR Pfam: PF00608; adeno\_fiber2; 9.  
 DR PRINTS: PRO0307; ADENOVSFIBRE.  
 KM Fiber protein.  
 SQ SEQUENCE 547 AA; 59113 MW; D1DDC36C4C935DE1 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 547;  
 Best Local Similarity 87.5%; Pred. No. 54;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AVALTSP 10  
 DB 201 AVALTSP 208

RESULT 24  
 ID ATKA ANASL STANDARD; PRT; 561 AA.

AC O9REX2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Potassium-translocating ATPase A chain (EC 3.6.3.12) (potassium-  
 DE translocating ATPase A chain) (ATP phosphohydrolase [potassium-  
 DE transporting] A chain) (potassium binding and translocating subunit  
 DE A).  
 GN KDPa.  
 OS Anabaena sp. (strain L31).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.

OX NCBI\_TaxID=29412;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Ballal A.D.; Gassel M.; Schlessinger E.; Rajaram H.; Apre S.K.;  
 RA Altendorf K.;

RT "The kdp operon of the heterocystous nitrogen-fixing cyanobacterium  
 RT Anabaena sp. strain L-31."  
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

CC - FUNCTION: ONE OF THE COMPONENTS OF THE HIGH-AFFINITY ATP-DRIVEN  
 CC POTASSIUM TRANSPORT (OR KDP) SYSTEM, WHICH CATALYZES THE  
 CC HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF HYDROGEN AND

CC POTASSIUM IONS (BY SIMILARITY).  
 CC - CATALYTIC ACTIVITY: ATP + H(2)O + K(+) (Out) = ADP + phosphate +  
 CC K(+) (In).

CC - SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Probable).  
 CC - SIMILARITY: BELONGS TO THE KDPa FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF213466; AAF19986.1; -;  
 DR InterPro: IPR004623; K\_ATPaseA.  
 DR TIGRPFAM: TIGR00680; kdpA; 1.

KM Hydrolyase; Transport; Potassium transport; Transmembrane;  
 KM Inner membrane.

FT TRANSMEM 2 22 POTENTIAL.  
 FT TRANSMEM 65 85 POTENTIAL.  
 FT TRANSMEM 135 155 POTENTIAL.  
 FT TRANSMEM 177 197 POTENTIAL.  
 FT TRANSMEM 253 273 POTENTIAL.  
 FT TRANSMEM 280 300 POTENTIAL.  
 FT TRANSMEM 327 347 POTENTIAL.  
 FT TRANSMEM 353 373 POTENTIAL.  
 FT TRANSMEM 378 398 POTENTIAL.  
 FT TRANSMEM 413 433 POTENTIAL.  
 FT TRANSMEM 482 502 POTENTIAL.  
 FT TRANSMEM 532 552 POTENTIAL.  
 SQ SEQUENCE 561 AA; 60776 MW; C8EE833F0E07D81 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 561;  
 Best Local Similarity 75.0%; Pred. No. 55;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSVAITYSP 8  
 DB 430 PSVAITYSP 437

RESULT 25  
 ID PTSA STRMU STANDARD; PRT; 664 AA.  
 AC P12655;  
 DT 01-OCT-1989 (Rel. 12, Created)

